Validated By CRFValidator v 1.0.3

Application No: 10537075 Version No: 2.0

Input Set:

Output Set:

Started: 2007-09-20 12:06:58.250
Finished: 2007-09-20 12:06:59.574
Elapsed: 0 hr(s) 0 min(s) 1 sec(s) 324 ms
Total Warnings: 5
Total Errors: 0
No. of SeqIDs Defined: 24
Actual SeqID Count: 24

Error code		Error Description
W	213	Artificial or Unknown found in <213> in SEQ ID (20)
W	213	Artificial or Unknown found in <213> in SEQ ID (21)
W	213	Artificial or Unknown found in <213> in SEQ ID (22)
W	213	Artificial or Unknown found in <213> in SEQ ID (23)
W	213	Artificial or Unknown found in <213> in SEQ ID (24)

SEQUENCE LISTING

<110> Kesseler, Maria
Zelinski, Thomas
Hauer, Bernhard

<120> L-RHAMNOSE-INDUCIBLE EXPRESSION SYSTEMS

<130> 12810-00091-US

<140> 10537075
<141> 2005-06-01

<150> PCT/EP2003/013367
<151> 2003-11-27

<150> DE 102 56 381.0
<151> 2002-12-02

<160> 24

<170> PatentIn version 3.3

<210> 1
<211> 2046
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (288)..(1121)
<223> coding for rhaS (positive regulator of rhaBAD operon)

<220>
<221> misc_feature
<222> (1108)..(2043)
<223> coding for rhaR (positive regulator of rhaRS operon)

<220>
<221> protein_bind
<222> (56)..(72)
<223> potential RhaS binding site

<220>
<221> protein_bind
<222> (89)..(105)
<223> potential RhaS binding site

<220>
<221> protein_bind
<222> (172)..(203)
<223> potential RhaR binding site

<220>
<221> protein_bind
<222> (210)..(241)
<223> potential RhaR binding site

<220>
<221> misc_feature
<222> (24)
<223> potential start of transcription (complement)

<400> 1
aatgtgatcc tgctgaattt cattacgacc agtctaaaaa gcgcctgaat tcgcgacctt 60
ctcgttactg acaggaaaat gggccattgg caaccaggga aagatgaacg tgatgatgtt 120
cacaatttgc tgaattgtgg tgatgtgatg ctcaccgcat ttcctgaaaa ttcacgctgt 180
atcttgaaaa atcgacgttt tttacgtggt tttccgtcga aaatttaagg taagaacctg 240
acctcgtgat tactatttcg ccgtgttgac gacatcagga ggccagtatg accgtattac 300
atagtgtgga tttttttccg tctggtaacg cgtccgtggc gatagaaccc cggctcccgc 360
aggcggattt tcctgaacat catcatgatt ttcatgaaat tgtgattgtc gaacatggca 420
cgggtattca tgtgtttaat gggcagccct ataccatcac cggtggcacg gtctgtttcg 480
tacgcgatca tgatcggcat ctgtatgaac ataccgataa tctgtgtctg accaatgtgc 540
tgtatcgctc gccggatcga tttcagtttc tcgccgggct gaatcagttg ctgccacaag 600
agctggatgg gcagtatccg tctcactggc gcgttaacca cagcgtattg cagcaggtgc 660
gacagctggt tgcacagatg gaacagcagg aaggggaaaa tgatttaccc tcgaccgcca 720
gtcgcgagat cttgtttatg caattactgc tcttgctgcg taaaagcagt ttgcaggaga 780
acctggaaaa cagcgcatca cgtctcaact tgcttctggc ctggctggag gaccattttg 840
ccgatgaggt gaattgggat gccgtggcgg atcaattttc tctttcactg cgtacgctac 900
atcggcagct taagcagcaa acgggactga cgcctcagcg atacctgaac cgcctgcgac 960
tgatgaaagc ccgacatctg ctacgccaca gcgaggccag cgttactgac atcgcctatc 1020
gctgtggatt cagcgacagt aaccactttt cgacgctttt tcgccgagag tttaactggt 1080
caccgcgtga tattcgccag ggacgggatg gctttctgca ataacgcgaa tcttctcaac 1140
gtatttgtac gccatattgc gaataatcaa cttcgttctc tggccgaggt agccacggtg 1200
gcgcatcagt taaaacttct caaagatgat ttttttgcca gcgaccagca ggcagtcgct 1260
gtggctgacc gttatccgca agatgtcttt gctgaacata cacatgattt ttgtgagctg 1320
gtgattgtct ggcgcggtaa tggcctgcat gtactcaacg atcgccctta tcgcattacc 1380
cgtggcgatc tcttttacat tcatgctgac gataaacact cctacgcttc cgttaacgat 1440
ctggttttgc agaatattat ttattgcccg gagcgtctga agctgaatct tgactggcag 1500
ggggcgattc cgggatttaa cgccagcgca gggcaaccac actggcgctt aggtagcatg 1560
gggatggcgc aggcgcggca ggttatcggt cagcttgagc atgaaagtag tcagcatgtg 1620
ccgtttgcta acgaaatggc tgagttgctg ttcgggcagt tggtgatgtt gctgaatcgc 1680
catcgttaca ccagtgattc gttgccgcca acatccagcg aaacgttgct ggataagctg 1740
attacccggc tggcggctag cctgaaaagt ccctttgcgc tggataaatt ttgtgatgag 1800
gcatcgtgca gtgagcgcgt tttgcgtcag caatttcgcc agcagactgg aatgaccatc 1860
aatcaatatc tgcgacaggt cagagtgtgt catgcgcaat atcttctcca gcatagccgc 1920
ctgttaatca gtgatatttc gaccgaatgt ggctttgaag atagtaacta tttttcggtg 1980
gtgtttaccc gggaaaccgg gatgacgccc agccagtggc gtcatctcaa ttcgcagaaa 2040
gattaa 2046

<210> 2
<211> 287
<212> DNA
<213> Escherichia coli

<220>
<221> promoter
<222> (1)..(287)
<223> rhaBAD promoter fragment containing rhaS and rhaR binding sites

<400> 2
actggcctcc tgatgtcgtc aacacggcga aatagtaatc acgaggtcag gttcttacct 60
taaattttcg acggaaaacc acgtaaaaaa cgtcgatttt tcaagataca gcgtgaattt 120

tcaggaaatg cggtgagcat cacatcacca caattcagca aattgtgaac atcatcacgt 180
tcatctttcc ctggttgcca atggcccatt ttcctgtcag taacgagaag gtcgcgaatt 240
caggcgcttt ttagactggt cgtaatgaaa ttcagcagga tcacatt 287

<210> 3
<211> 125
<212> DNA
<213> Escherichia coli

<220>
<221> promoter
<222> (1)..(125)
<223> rhaBAD promoter fragment containing RhaS binding site

<400> 3
ttgtgaacat catcacgttc atctttccct ggttgccaat ggcccatttt cctgtcagta 60
acgagaaggt cgcgaattca ggcgcttttt agactggtcg taatgaaatt cagcaggatc 120
acatt 125

<210> 4
<211> 123
<212> DNA
<213> Escherichia coli

<220>
<221> promoter
<222> (1)..(123)
<223> rhaBAD promoter fragment containing RhaS binding site

<400> 4
atcaccacaa ttcagcaaat tgtgaacatc atcacgttca tctttccctg gttgccaatg 60
gcccattttc ctgtcagtaa cgagaaggtc gcgaattcag gcgcttttta gactggtcgt 120
aat 123

<210> 5
<211> 51
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)..(51)
<223> palindromic RhaS binding site of rhaBAD promoter

<400> 5
atctttccct ggttgccaat ggcccatttt cctgtcagta acgagaaggt c 51

<210> 6
<211> 1071
<212> DNA
<213> Alcaligenes faecalis

<220>

<221> CDS
<222> (1)..(1068)
<223> coding for nitrilase

<400> 6

```
atg cag aca aga aaa atc gtc cgg gca gcc gcc gta cag gcc gcc tct   48
Met Gln Thr Arg Lys Ile Val Arg Ala Ala Ala Val Gln Ala Ala Ser
  1               5                  10                  15
ccc aac tac gat ctg gca acg ggt gtt gat aaa acc att gag ctg gct   96
Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu Ala
             20                  25                  30
cgt cag gcc cgc gat gag ggc tgt gac ctg atc gtg ttt ggt gaa acc   144
Arg Gln Ala Arg Asp Glu Gly Cys Asp Leu Ile Val Phe Gly Glu Thr
         35                  40                  45
tgg ctg ccc gga tat ccc ttc cac gtc tgg ctg ggc gca ccg gcc tgg   192
Trp Leu Pro Gly Tyr Pro Phe His Val Trp Leu Gly Ala Pro Ala Trp
     50                  55                  60
tcg ctg aaa tac agt gcc cgc tac tat gcc aac tcg ctc tcg ctg gac   240
Ser Leu Lys Tyr Ser Ala Arg Tyr Tyr Ala Asn Ser Leu Ser Leu Asp
 65                  70                  75                  80
agt gca gag ttt caa cgc att gcc cag gcc gca cgg acc ttg ggt att   288
Ser Ala Glu Phe Gln Arg Ile Ala Gln Ala Ala Arg Thr Leu Gly Ile
                 85                  90                  95
ttc atc gca ctg ggt tat agc gag cgc agc ggc ggc agc ctt tac ctg   336
Phe Ile Ala Leu Gly Tyr Ser Glu Arg Ser Gly Gly Ser Leu Tyr Leu
            100                 105                 110
ggc caa tgc ctg atc gac gac aag ggc gag atg ctg tgg tcg cgt cgc   384
Gly Gln Cys Leu Ile Asp Asp Lys Gly Glu Met Leu Trp Ser Arg Arg
        115                 120                 125
aaa ctc aaa ccc acg cat gta gag cgc acc gta ttt ggt gaa ggt tat   432
Lys Leu Lys Pro Thr His Val Glu Arg Thr Val Phe Gly Glu Gly Tyr
    130                 135                 140
gcc cgt gat ctg att gtg tcc gac aca gaa ctg gga cgc gtc ggt gct   480
Ala Arg Asp Leu Ile Val Ser Asp Thr Glu Leu Gly Arg Val Gly Ala
145                 150                 155                 160
cta tgc tgc tgg gag cat ttg tcg ccc ttg agc aag tac gcg ctg tac   528
Leu Cys Cys Trp Glu His Leu Ser Pro Leu Ser Lys Tyr Ala Leu Tyr
                165                 170                 175
tcc cag cat gaa gcc att cac att gct gcc tgg ccg tcg ttt tcg cta   576
Ser Gln His Glu Ala Ile His Ile Ala Ala Trp Pro Ser Phe Ser Leu
            180                 185                 190
tac agc gaa cag gcc cac gcc ctc agt gcc aag gtg aac atg gct gcc   624
Tyr Ser Glu Gln Ala His Ala Leu Ser Ala Lys Val Asn Met Ala Ala
        195                 200                 205
tcg caa atc tat tcg gtt gaa ggc cag tgc ttt acc atc gcc gcc agc   672
Ser Gln Ile Tyr Ser Val Glu Gly Gln Cys Phe Thr Ile Ala Ala Ser
    210                 215                 220
agt gtg gtc acc caa gag acg cta gac atg ctg gaa gtg ggt gaa cac   720
Ser Val Val Thr Gln Glu Thr Leu Asp Met Leu Glu Val Gly Glu His
225                 230                 235                 240
aac gcc ccc ttg ctg aaa gtg ggc ggc ggc agt tcc atg att ttt gcg   768
Asn Ala Pro Leu Leu Lys Val Gly Gly Gly Ser Ser Met Ile Phe Ala
                245                 250                 255
ccg gac gga cgc aca ctg gct ccc tac ctg cct cac gat gcc gag ggc   816
Pro Asp Gly Arg Thr Leu Ala Pro Tyr Leu Pro His Asp Ala Glu Gly
            260                 265                 270
ttg atc att gcc gat ctg aat atg gag gag att gcc ttc gcc aaa gcg   864
```

```
Leu Ile Ile Ala Asp Leu Asn Met Glu Glu Ile Ala Phe Ala Lys Ala
        275                 280                 285
atc aat gac ccc gta ggc cac tat tcc aaa ccc gag gcc acc cgt ctg   912
Ile Asn Asp Pro Val Gly His Tyr Ser Lys Pro Glu Ala Thr Arg Leu
    290                 295                 300
gtg ctg gac ttg ggg cac cga gac ccc atg act cgg gtg cac tcc aaa   960
Val Leu Asp Leu Gly His Arg Asp Pro Met Thr Arg Val His Ser Lys
305                 310                 315                 320
agc gtg acc agg gaa gag gct ccc gag caa ggt gtg caa agc aag att   1008
Ser Val Thr Arg Glu Glu Ala Pro Glu Gln Gly Val Gln Ser Lys Ile
                325                 330                 335
gcc tca gtc gct atc agc cat cca cag gac tcg gac aca ctg cta gtg   1056
Ala Ser Val Ala Ile Ser His Pro Gln Asp Ser Asp Thr Leu Leu Val
            340                 345                 350
caa gag ccg tct tga                                             1071
Gln Glu Pro Ser
        355
```

<210> 7
<211> 356
<212> PRT
<213> Alcaligenes faecalis

<400> 7

```
Met Gln Thr Arg Lys Ile Val Arg Ala Ala Ala Val Gln Ala Ala Ser
  1               5                  10                  15
Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu Ala
             20                  25                  30
Arg Gln Ala Arg Asp Glu Gly Cys Asp Leu Ile Val Phe Gly Glu Thr
         35                  40                  45
Trp Leu Pro Gly Tyr Pro Phe His Val Trp Leu Gly Ala Pro Ala Trp
     50                  55                  60
Ser Leu Lys Tyr Ser Ala Arg Tyr Tyr Ala Asn Ser Leu Ser Leu Asp
 65                  70                  75                  80
Ser Ala Glu Phe Gln Arg Ile Ala Gln Ala Ala Arg Thr Leu Gly Ile
                 85                  90                  95
Phe Ile Ala Leu Gly Tyr Ser Glu Arg Ser Gly Gly Ser Leu Tyr Leu
            100                 105                 110
Gly Gln Cys Leu Ile Asp Asp Lys Gly Glu Met Leu Trp Ser Arg Arg
        115                 120                 125
Lys Leu Lys Pro Thr His Val Glu Arg Thr Val Phe Gly Glu Gly Tyr
    130                 135                 140
Ala Arg Asp Leu Ile Val Ser Asp Thr Glu Leu Gly Arg Val Gly Ala
145                 150                 155                 160
Leu Cys Cys Trp Glu His Leu Ser Pro Leu Ser Lys Tyr Ala Leu Tyr
                165                 170                 175
Ser Gln His Glu Ala Ile His Ile Ala Ala Trp Pro Ser Phe Ser Leu
            180                 185                 190
Tyr Ser Glu Gln Ala His Ala Leu Ser Ala Lys Val Asn Met Ala Ala
        195                 200                 205
Ser Gln Ile Tyr Ser Val Glu Gly Gln Cys Phe Thr Ile Ala Ala Ser
    210                 215                 220
Ser Val Val Thr Gln Glu Thr Leu Asp Met Leu Glu Val Gly Glu His
225                 230                 235                 240
Asn Ala Pro Leu Leu Lys Val Gly Gly Gly Ser Ser Met Ile Phe Ala
                245                 250                 255
```

```
Pro Asp Gly Arg Thr Leu Ala Pro Tyr Leu Pro His Asp Ala Glu Gly
            260                 265                 270
Leu Ile Ile Ala Asp Leu Asn Met Glu Glu Ile Ala Phe Ala Lys Ala
        275                 280                 285
Ile Asn Asp Pro Val Gly His Tyr Ser Lys Pro Glu Ala Thr Arg Leu
    290                 295                 300
Val Leu Asp Leu Gly His Arg Asp Pro Met Thr Arg Val His Ser Lys
305                 310                 315                 320
Ser Val Thr Arg Glu Glu Ala Pro Glu Gln Gly Val Gln Ser Lys Ile
                325                 330                 335
Ala Ser Val Ala Ile Ser His Pro Gln Asp Ser Asp Thr Leu Leu Val
            340                 345                 350
Gln Glu Pro Ser
        355
```

<210> 8
<211> 1260
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(1257)
<223> coding for rhaA (L-rhamnose isomerase)

<400> 8

```
atg acc act caa ctg gaa cag gcc tgg gag cta gcg aaa cag cgt ttc   48
Met Thr Thr Gln Leu Glu Gln Ala Trp Glu Leu Ala Lys Gln Arg Phe
  1               5                  10                  15
gcg gcg gtg ggg att gat gtc gag gag gcg ctg cgc caa ctt gat cgt   96
Ala Ala Val Gly Ile Asp Val Glu Glu Ala Leu Arg Gln Leu Asp Arg
             20                  25                  30
tta ccc gtt tca atg cac tgc tgg cag ggc gat gat gtt tcc ggt ttt   144
Leu Pro Val Ser Met His Cys Trp Gln Gly Asp Asp Val Ser Gly Phe
         35                  40                  45
gaa aac ccg gaa ggt tcg ctg acc ggg ggg att cag gcc aca ggc aat   192
Glu Asn Pro Glu Gly Ser Leu Thr Gly Gly Ile Gln Ala Thr Gly Asn
     50                  55                  60
tat ccg ggc aaa gcg cgt aat gcc agt gag cta cgt gcc gat ctg gaa   240
Tyr Pro Gly Lys Ala Arg Asn Ala Ser Glu Leu Arg Ala Asp Leu Glu
 65                  70                  75                  80
cag gct atg cgg ctg att ccg ggg ccg aaa cgg ctt aat tta cat gcc   288
Gln Ala Met Arg Leu Ile Pro Gly Pro Lys Arg Leu Asn Leu His Ala
                 85                  90                  95
atc tat ctg gaa tca gat acg cca gtc tcg cgc gac cag atc aaa cca   336
Ile Tyr Leu Glu Ser Asp Thr Pro Val Ser Arg Asp Gln Ile Lys Pro
            100                 105                 110
gag cac ttc aaa aac tgg gtt gaa tgg gcg aaa gcc aat cag ctc ggt   384
Glu His Phe Lys Asn Trp Val Glu Trp Ala Lys Ala Asn Gln Leu Gly
        115                 120                 125
ctg gat ttt aac ccc tcc tgc ttt tcg cat ccg cta agc gcc gat ggc   432
Leu Asp Phe Asn Pro Ser Cys Phe Ser His Pro Leu Ser Ala Asp Gly
    130                 135                 140
ttt acg ctt tcc cat gcc gac gac agc att cgc cag ttc tgg att gat   480
Phe Thr Leu Ser His Ala Asp Asp Ser Ile Arg Gln Phe Trp Ile Asp
145                 150                 155                 160
```

```
cac tgc aaa gcc agc cgt cgc gtt tcg gcc tat ttt ggc gag caa ctc   528
His Cys Lys Ala Ser Arg Arg Val Ser Ala Tyr Phe Gly Glu Gln Leu
                165                 170                 175
ggc aca cca tcg gtg atg aac atc tgg atc ccg gat ggt atg aaa gat   576
Gly Thr Pro Ser Val Met Asn Ile Trp Ile Pro Asp Gly Met Lys Asp
            180                 185                 190
atc acc gtt gac cgt ctc gcc ccg cgt cag cgt ctg ctg gca gca ctg   624
Ile Thr Val Asp Arg Leu Ala Pro Arg Gln Arg Leu Leu Ala Ala Leu
        195                 200                 205
gat gag gtg atc agc gag aag cta aac cct gcg cac cat atc gac gcc   672
Asp Glu Val Ile Ser Glu Lys Leu Asn Pro Ala His His Ile Asp Ala
    210                 215                 220
gtt gag agc aaa ttg ttt ggc att ggc gca gag agc tac acg gtt ggc   720
Val Glu Ser Lys Leu Phe Gly Ile Gly Ala Glu Ser Tyr Thr Val Gly
225                 230                 235                 240
tcc aat gag ttt tac atg ggg tat gcc acc agc cgc cag act gcg ctg   768
Ser Asn Glu Phe Tyr Met Gly Tyr Ala Thr Ser Arg Gln Thr Ala Leu
                245                 250                 255
tgc ctg gac gcc ggg cac ttc cac ccg act gaa gtg att tcc gac aag   816
Cys Leu Asp Ala Gly His Phe His Pro Thr Glu Val Ile Ser Asp Lys
            260                 265                 270
att tcc gcc gcc atg ctg tat gtg ccg cag ttg ctg ctg cac gtc agc   864
Ile Ser Ala Ala Met Leu Tyr Val Pro Gln Leu Leu Leu His Val Ser
        275                 280                 285
cgt ccg gtt cgc tgg gac agc gat cac gta gtg ctg ctg gat gat gaa   912
Arg Pro Val Arg Trp Asp Ser Asp His Val Val Leu Leu Asp Asp Glu
    290                 295                 300
acc cag gca att gcc agt gag att gtg cgt cac gat ctg ttt gac cgg   960
Thr Gln Ala Ile Ala Ser Glu Ile Val Arg His Asp Leu Phe Asp Arg
305                 310                 315                 320
gtg cat atc ggc ctt gac ttc ttc gat gcc tct atc aac cgc att gcc   1008
Val His Ile Gly Leu Asp Phe Phe Asp Ala Ser Ile Asn Arg Ile Ala
                325                 330                 335
gcg tgg gtc att ggt aca cgc aat atg aaa aaa gcc ctg ctg cgt gcg   1056
Ala Trp Val Ile Gly Thr Arg Asn Met Lys Lys Ala Leu Leu Arg Ala
            340
```